

White Paper Report

Report ID: 98493

Application Number: HD5056409

Project Director: Hanna Griff-Sleven (hgriff@eldridgestreet.org)

Institution: Eldridge Street Project, Inc./Museum at Eldridge Street

Reporting Period: 4/1/2009-3/31/2010

Report Due: 6/30/2010

Date Submitted: 3/24/2010

NEH Digital Humanities Start-Up

Grant HD-50564-09

**ILLUMINATING THE IMMIGRANT EXPERIENCE: LEVEL I
WHITE PAPER**

ELDRIDGE STREET PROJECT, INC. (aka, Museum at Eldridge Street)

12 Eldridge Street

New York, NY 10002

www.eldridgestreet.org

Project Director, Hanna Griff-Sleven hgriff@eldridgestreet.org

Institutional Grant Administrator, Eva Bruné ebrune@eldridgestreet.org

Executive Director, Bonnie Dimun bdimun@eldridgestreet.org

A Digital Humanities Start-Up Project funded by the NEH in partnership with the IMLS.

Work completed March 23, 2010

Table of Contents

Statement of Project’s Initial Goals and Objectives	3
Project Participants	3-4
Methodology and Activities	
Statement of problem	4
Methodology we employed to answer the problem	4-5
Audience for this project	5-6
Summary of Existing Tools	6-8
The Seven Entities & ‘Producer-Users’	
The Seven Entities	8-10
Links Between Entities	10
Custom Entities	11
Eldridge Street Sample Profiles	
From Concept to Practice	11-12
Example of One Profile	12
Final Meeting Discussion	12
Open Systems	13
Content and Form	13
Relationships	13-14
Next Steps	14

Statement of Project's Initial Goals and Objectives

NEH funding of \$24,382 was provided to the Museum at Eldridge Street (legal name: Eldridge Street Project, Inc.) to support a series of planning meetings with a team of humanities scholars, K-12 teachers, and interactive technology experts to explore how separate databases and historical resources could be brought together to contextualize and share a rich and complex picture of a community.

By focusing on a small sampling of original congregants, we sought to investigate the occupational trajectories, family economies, and geographic and economic mobility of these Eastern European Jewish immigrants who made the Lower East Side the largest Jewish community in the world at the turn of the 20th century. We sought to move beyond the current web-based offerings of individual web-sites containing historical data, seeking to conceptualize the software and the data structure necessary to “contextualize the historical information” we gathered from our on-site archives and from the websites and archives of our project partners.

Beginning in April 2009, we convened our team to plan for the development of a highly interactive physical installation (what we call a “history table,” based on the two AAM award-winning history tables designed by the Museum at Eldridge Street and Potion in 2006) and web-based tool to be used at the Museum at Eldridge Street. The new software concept would be designed to enable the Museum to create its interactive resources, and would make these resources available to individuals and organizations seeking to create their own community profiles. Our second goal was to create a plan for open source software that uses our findings and the lessons learned in this process to allow other communities and organizations nationwide, and of all cultural backgrounds, to engage in a similar process.

These goals were realized. At the same time, the project yielded other unexpected outcomes: the Museum at Eldridge Street, footnotes.com and Potion are exploring how we could work together on other projects of mutual benefit, and strong relationships were advanced between footnotes.com and the New York Public Library. We are grateful to the NEH for enabling these developments.

Project Participants

Project participants included representatives of three technology firms to lead the software development process and potential front-end considerations (*Potion*, represented by **Philip Tionsgen** and **Jared Schiffman**, with guests **Caroline Brown** and **Miranda Heineman** joining the team; *footnotes.com*, represented by **Roger Bell**, with guest **Chris Willis** joining the project for the second and third meetings; and the *New York City Public Library*, represented by **Joshua Greenberg**, whose background as an historian and technology expert proved particularly valuable; two K-12 teachers to lead discussions about the ways K-12 teachers and lay people might access and use the digital resource in their classroom and for historical research (**Carol Baldridge** and **William Ebert**); humanities researchers who shared their expertise about their resources and database issues (**Karen Franklin**, genealogist, and guest **Dr. Susan Malbin** from the American Jewish Historical Society who represented executive director **Evan Kingsley** (who attended the first meeting) at the second and third meetings); humanities scholars (**Dr. Thomas Kessner** and **Dr. Jeffrey Gurock**); and Museum at Eldridge Street staff (**Dr. Hanna Griff-Sleven**, Director for Programs and the Family History Center at the Museum at Eldridge Street, project director; and guests **Dr. Bonnie Dimun**, executive director, **Nina Cohen**, Education and Special Projects Associate, and **Eva Bruné**, Vice President for Institutional Advancement). Their profiles are included as *Addendix I: Participant Profiles*.

Participants attended three, six-hour meetings that took place at the Museum at Eldridge Street. Between meetings, team and guest members contributed thoughts and comments via an on-line “base camp” that Potion set up for the project. Separate telephone meetings were held with different sub-groups of team members and guests. In the course of these dialogues, we learned much about each other’s organizational projects, leading to tangential discussions. Between the second and third meetings, the team members focused on ‘next steps’ for development

of the software and the Museum's preparations for the 'front-end,' whether it be another history table, more expanded and interactive website, and/or educational adjunct devices and/or activities.

Methodology and Activities

Statement of Problem

From the beginning, the team sought to create a model that could be used by the Museum at Eldridge Street, as well as be replicated in any other American city, using any historical site, church, school, or other location, as an anchor to collect information about individuals and families. This information could be used to connect the trajectory of each family to the broader community. Humanities issues were inextricably linked to the central technological issues faced by our team. *How do you encode a sense of the process into the database? How can we provide that information in our database/exhibit in a smooth and visually exciting manner?*

Methodology we employed to address the Problems

We approached the project through a rich humanities-based discourse, considering a variety of potential organizing themes and subjects: immigrant life, geography, our site's congregational life as it connected to and helped shape the broader community, and the broad subject of 'livelihood,' or other potential organizing principles.

We discussed the varied immigrant experience, as well as the specific experience of turn-of-the-century East European Jewish immigration to the Lower East Side: the roles of economic mobility and education; issues of cultural change over time, including migration patterns once they arrived in America, as well as occupational mobility and the ways in which the immigrants adapted to their new communities.

How could the software we were hoping to develop trace individual identities? How could the software also trace ways in which different communities intersected and overlapped over generations in our Lower East Side neighborhood? And, how could the software be able to reflect data and stories in a way that enables other organizations engaged in studying community?

We segued into a discussion of the congregants of the Eldridge Street's historic site, as this was the population that Eldridge Street would use to create its database and project. In addition to looking at them as individuals, we discussed ways in which data could be collected and the software used to reflect upon individual and collective influence on the Eldridge Street Synagogue's traditions and practices – the Americanization of Jewish practice as reflected through the lens of the Eldridge Street Synagogue. As Dr. Kessner reflected: How can one congregation's experience illuminate a broader community's history? Did the synagogue community partner with other ethnic groups? How did the synagogue community adapt to the changing face and nature of the neighborhood of which they were part? The data and stories, oral histories and photographs we collect should reflect the ways in which these immigrants assimilated and adapted to their new environment.

How would we identify the change-makers? Who sought to preserve the East European "old ways?" Was there a way to correlate levels of affluence with a desire for change? Or was there a desire for change in Jewish practice as a way to keep the next generation connected to religious life? The methods and findings derived from looking at the Eldridge Street Synagogue's early congregants should be applicable and comparable to other synagogue (and non-Jewish) experiences in the U.S. By collecting stories and data relating to religious practice in communities, we could further explore these questions.

We then explored the meaning and definition of 'community,' looking at the nature of community infrastructure and, in particular, the social welfare web they constructed to support their unemployed brethren, to bury the poor, offer advice and assistance to newcomers making their way in American society. We also examined the East European Jewish immigrants' relationship with other immigrant groups settling on the Lower East Side during the same period, considering bias and discrimination that each community faced in terms of K-

12 and advanced education, access to jobs and social clubs, and in day-to-day life, including such issues as local and national political power, the notion of bloc voting, and political activism.

Our conversation shifted to the broader neighborhood and the Lower East Side's geography as a potential organizing factor. We quickly realized that geography had a host of challenges: canals became streets, street names changed or were reshaped to fit a changing landscape as the section grew over time. And the neighborhood dynamics also changed over time. These issues would have to be addressed in the development of the software.

Our discussion also explored "livelihood" as an organizing principle and entry point. Livelihood – the struggle to build businesses, raise families, and otherwise make ends meet is a concept to which all people can relate. As we explored our project from the thematic approach of 'livelihood,' we moved from the individual story to the community story and back again:

Individual occupational profiles: What traditions of work, skill, and aspirations did the immigrants bring with them? In what businesses and trades did these 19th century immigrants engage? How long did it take for newcomers to advance beyond their original low working class roots to become part of an entrepreneurial class? What was the work experience of the immigrants' American-born children? To what extent was it a one generational phenomenon? Or did it change from the first to second generation? The data and stories would be gathered to help explore and possibly (at least from the Eldridge Street Synagogue's congregants stories) answer some of these questions.

Gender and Age Factors: What were the gender differences in occupational development? Was there more gender equality or less in the ethnic community? And what about the less-told story of the economic and social adjustments of those who came here as pre-teens and young adults (ages 6-14 or so). Were their experiences, arguably, in terms of work, education and adjustment to America different both from their parents and their second generation (American-born) siblings?

Community occupational profile: What were the industrial/commercial bases of the local economy? What was the occupational setting into which the immigrants came when they arrived? How did this change over the years? How was it affected by the Depression? Did people move away from the community primarily to improve their economic situation, in search of better and healthier housing, or a desire to break with the first generation culture of the ethnic community?

By the close of our second meeting, and using the Eldridge Street Synagogue's congregants and their early 20th century Lower East Side neighborhood as a model, we discussed the many reasons why the great wave of East European Jewry settled on the Lower East Side came to the United States at the end of the 19th century. Two important reasons – to seek freedom of worship and economic opportunity – interested the team. Certainly, the immigrants established businesses, trade networks, unions, fraternal societies, houses of worship and other communities to help them navigate a new country and whose legacies live on today. We agreed that many immigrant groups seek the opportunity to pursue livelihoods to build their new lives. We felt that Museum visitors of all backgrounds would appreciate 'livelihood' as an organizing principle, drawing parallels to their own families' settlement in America. However, we recognized that "livelihood" was just one of many – not the only – option that could serve as an organizing principle for the software.

Audience for this Project

Throughout our discussions during the first and second meetings, we struggled to structure the project to simultaneously meet the needs of the Museum at Eldridge Street and an amorphous 'everyman end user.' At times, we were hand-tied by wanting the software to be all things to all people.

In exploring the end-user, we tried to determine who the ultimate audience, beyond Eldridge Street's visitors, were for the open-source software. Was the audience an individual end-user sitting at a computer in the

comfort of their own home, building a database based on individuals from their church or neighborhood? Is the software for individuals at home? What is the difference between scholarly research and the ‘everyman/woman public’ research? Is it for educators looking to create an ongoing local history project? Or is it for organizations like the Museum at Eldridge Street, e.g. local historical societies, churches, and schools around the country? We knew that the Museum at Eldridge Street was the model for the audience, in that it is an institution with local scholars who are doing research and need a place to store their data. Could we focus on this audience for this software? If MAES is a producer of content should the primary audience of this software be other producers?

We decided to focus our discussion on Eldridge Street’s needs as primary and defined other end-users of the software as ‘producers’ – organizations like Eldridge Street – museums, historical societies, churches who would build the database and create a ‘front-end’ that visitors would then view (rather than individuals using the software for their own personal projects). In defining our end users as ‘producers,’ we were able to advance to what the data needed to include and, eventually, how the ‘front end’ might work.

At this point, Potion, footnote.com, and the NYPL representatives took the team on a thorough review of existing tools to solve our project’s challenges. How did they succeed and where did they fail? What were the rights and ownership issues involved? What could we learn or leverage from existing tools for application to our software?

Summary of Existing Tools

There are no lack of database tools and websites that deal with the type of content that this whitepaper addresses. Local history is well represented on the internet in a variety of forms, from the websites of historical societies and museums to the personal blogs of amateur historians. What is difficult to find, however, are well considered systems for gathering and storing local history in an organized manner. The gathering and storing of information is not a challenge. It is the organization that is the difficult part, and it is the organization that ultimately determines how the data can be used and re-used.

Unfortunately, most of the historical information stored on the web today is either unformatted or is formatted inconsistently. While this doesn’t bother most web consumers, it makes the information difficult to re-present in other formats and makes it challenging to search and process in a meaningful way. Nonetheless, there are examples of online information sets that are properly formatted (also referred to as “structured”), but unfortunately, they do not capture the domain of local history particularly well, or in a way that would be useful to a small institution such as Eldridge Street. The database schema proposed in this whitepaper hopes to provide a structured system for local history information storage that will work for small to medium-scale institutions. Before proceeding with the proposed system, it is worth considering the systems that are available today that exists in and around this domain.

Traditional Databases. Museums and historical institutions are common customers of database companies. Many use “The Museum System” to manage their collection or FileMakerPro to manage their content. These systems work very well under very specific conditions, namely, when the data is already highly formatted. Unfortunately, the stories of local history are not so neat. They resist being contained and are rich with cross-connections or links, something that traditional database do not handle well. They are generally organized around objects, or documents or single events in time, not around the people, places and political and social movements of local history. While all of the systems presented here are grounded in traditional databases, each has certain unique characteristics that allow them to transcend the formulaic tabular structure that most common databases imply.

Wikipedia. Wikipedia is by far the largest collection of user-generated historical information on the internet. Wikipedia is one instance of a particular technology, simply called “wiki” that enables groups of people to edit articles online and link them to one another. One may ask, and rightly so, “Why not use Wikipedia as a storage venue for Eldridge Street’s local history?” Certainly, it is capable of storing the information, it is easy to

use and has the added benefit of tens of thousands of Wikipedia contributors around the world. Unfortunately, the data stored on Wikipedia is not structured in a way that makes it amenable to re-use in different formats. Furthermore, it lacks the semantic underpinnings necessary to process its own data in a meaningful way. For example, if one searches Wikipedia for the last name “Boxer”, it will likely bring up every boxer that ever has ever boxed, since it doesn’t differentiate between people and professions. Furthermore, if you want a quick answer as to the year that an event occurred, it will not be able to help you because it has no sense of dates, more so than it does prices or other numbers. These examples are the tip of the iceberg in terms of the type of searching and processing that Wikipedia cannot handle. Furthermore, Wikipedia cannot handle multiple versions of a single story, which as any historian knows, is essential to a complete understanding. It is based around a model of consensus, not completeness. Last, a totally public and open data system, there’s no guarantee to the lifespan of the data. If an exhibit were linked to a specific article, and that article disappears, it is difficult to recover and will break any system that depends on it. All of this being said, Wikipedia is a wonderful system with a vast collection of knowledge, which is unfortunately, just not right for the needs of Eldridge Street.

Footnote.com. Footnote.com is an online database that is centered on historical documents. Over the past five years, Footnote has amassed an incredible collection of almost 70 million documents, mostly scanned from government archives and also from historical institutions. Footnote not only scans the documents but processes them as well, so that every scanned word on page is also represented as digital text. This makes Footnote a truly impressive research tool, since one can now digitally search enormous documents like the US 1930 Census, which previously had to be searched by hand. In addition to this functionality, Footnote allows its users to annotate the documents and even small sections within the documents with their own historical information. Specifically, there are four categories of annotation: people, places, dates and other. Through these annotations, the documents take on meaningful historical context and detail, and turn the document into a living reference point in history. While Footnote could certainly be a superb resource for Eldridge Street (especially for the Census and newspaper data), it is not a natural fit for the recording of local histories, most of which do not have strict documentary support. More to the point, finding a document for every story to be told would be a difficult task. In general, the stories of Eldridge Street are centered on people and places and events, while the media plays a supporting rather than a starring role.

Omeka. Omeka is an online web publishing tool developed by the Center for History and New Media at George Mason University. It is specifically designed for use by museums and historical societies to create online versions of their exhibitions. It is also used by scholars and students within educational institutions to create collections of media and research around areas of interest. Omeka’s raison d’être is about making web publishing simple and easy, and in this domain, it truly succeeds. It is also intent on being a good citizen of the internet, conforming to open standards and allowing content to be re-purposed more easily. What is nice about Omeka is that its websites still maintain their own look and feel (see <http://omeka.org/showcase/> for a large selection of Omeka sites). It is certainly unlike Wikipedia where every article looks like all the others. Eldridge Street could definitely use Omeka to create wonderful websites around the stories it has to tell. In many ways Omeka provides much more than Eldridge Street needs, and in some ways it provides less. As discussed earlier in this document, Eldridge needs a database solution that will allow for growth over many years and allow for the data to be re-presented in different media formats, both on the web and via installed interactive tables. While Omeka certainly contains a database, it is purposefully abstracted to make the overall system easier to use. Furthermore, Omeka’s database is specifically designed for presentation by Omeka’s web layer. Utilizing the data contained in Omeka’s database with other media systems might be difficult. Finally, the database contained with Omeka is not intended to be used as a research repository. It exists to store content that is ready for presentation, not necessarily information that exists solely for research purposes. All of these points notwithstanding, Omeka could be a great tool for Eldridge to use if and when it decides to create outward facing websites for its local history content.

HEML. HEML stands for Historical Event Markup Language. It is not a widely used system and is not even a database, but it is especially relevant to this whitepaper and to the needs of Eldridge Street. HEML is an open-source, text-based format for the representation of historic events in a standardized fashion. HEML derives from XML, which, like HTML, is a standardized text-based format for the exchange of information. While XML

is truly generic, HEML has been designed specifically represent historic events. An historic event within HEML is defined by the following elements (source: <http://heml.mta.ca/heml-cocoon/description>):

1. A label to name the event
2. One or more keywords that group conceptually similar events
3. A location in which the event took place
4. A 'chronology' of the event, describing the time in or at which it event took place
5. A list of persons or groups of persons who participate in the event
6. A list of evidence for the event, either in physical form (such as printed books or 16mm film) or as a web resource

As far as we know, this is the only open standard in existence that tries to capture historic event information in this way. HEML is well considered and has been peer-reviewed in academic journals. In many ways, HEML forms the underpinnings of the system being proposed within this whitepaper. Unfortunately, since HEML is simply a text format, and not a database or piece of software, it does not provide the necessary functionality required by Eldridge Street or similar institutions. As a minor point of criticism, we feel that HEML places too much emphasis on events (focusing on ultimate presentation via a timeline) and not enough emphasis on the people or places involved in those events. Also, it provides little opportunity for the inclusion of higher-level concepts such as political movements or migrations of the population. Nonetheless, HEML is a major innovation in the field of digital history and one that the proposed system owes homage.

The Seven Entities & 'Producer-Users'

The Seven Entities

We have defined seven entities that together comprise the complete schema for the database. These entities are high-level “objects” that are modeled after real-world entities, and should be relatively clear to non-technical users of the database. Each entity is defined by a small set of properties, also referred to as “fields” in database parlance. Again, the minimal set of properties was chosen for each entity for the reasons mentioned above. The seven entities are presented in the first column of the following table, with their respective properties displayed in the second column. Examples of each property are shown in the third column.

It is worth discussing briefly how each entity is intended to be used in practice. This information should be included in the database entry interface as well, so that all users inputting information understand how best to enter it.

See next page for diagram.

Entity	Properties	Example
Person	a. first name b. middle name c. last name d. occupations e. description	a. Isaac b. c. Gellis d. Owner of Isaac Gellis' Delicatesson... e.
Event	a. name b. start date day c. start date month d. start date year e. end date day f. end date month g. end date year h. description	a. Kosher Meat Boycott b. c. April d. 1902 e. f. May g. 1902 h. The Kosher Meat Boycott began when...
Location	a. name b. address c. street d. city e. state f. country g. description	a. Eldridge Street Synagogue b. 12 c. Eldridge Street d. New York e. New York f. USA g. Eldridge Street Synagogue was built...
Object	a. name b. description	a. Seat #20 b. This seat in the balcony of...
Institution	a. name b. description	a. Eldridge Street Board Members b. The board of Eldridge Street was...
Theme	a. name b. description	a. Role of women b. The role of women cannot be underestimated in the...
Media	a. name b. image c. audio d. movie	a. photo of Eldridge Street Synagogue b. c. d.
Link	a. description	a. John and Sam are brothers

Person. This is meant to represent actual individuals, either alive or dead. First and last name are required. Middle name is optional. The description, which is also optional, can be a phrase or single sentence or could be a lengthy description. The description may take the form of a short biography, though should avoid inclusion of lengthy description of events that involve other people, since these should be captured in ‘event’ entities.

Event. An event, in its most basic form, is a date and a name associated with that date. The event may also optionally have an end date to indicate a span of time. It is important to understand that events may span a great duration, during which time other related events may occur. For example, “World War I” may exist as event entity spanning from 1914 to 1918, while specific battles within the war exist as their own events during that time. Since the precise day or even month of an event may not be known, these properties are optional. The description of the event may be of any length and should describe the happenings of that event, but avoid detailed descriptions of the location or people involved, since this information can be captured in other entities.

Location. A location is simply an address and a name associated with that address. Since certain locations (such as “the Smith’s corn field”) may not have an exact street address, a name alone may suffice. Ideally, the city, state and country should be part of every location. The description of the location may be of any length, and should focus on the physical properties of the location, with brief reference to the events that happened there and people who occupied the space. Again, description of the people and events should take place in their own entities.

Object. An object is represented by its name and a description. The description of an object, much like locations, should focus on its physical properties and its relevance to individual people and events without describing either of these entities in full. If the database is continually referencing a specific class objects, such as menorahs, it may be worth considering adding administrator-defined properties to the entity. This process is described in more detail in the “Custom Properties” section that follows.

Institution. An institution is a group of people who have organized themselves in some way. It could be a company, a social group, a political party or a book club. It is assumed that the “person” entities for the members of the institution would necessarily be linked to the entity for the institution itself. The description of the institution may be of any length and should include discussion of the founding charter or reason for organization and may also include a brief history of the organization. Of course, the major events in the institution's history should be represented in the “event” entities that are linked to the institution.

Theme. Theme entities are the only entities in the schema that does not have singular real world counterparts. Theme entities are a place for historians and curators to link people, events and locations together and across time to tell a broader story. Themes may serve as umbrellas under which a collection of entities are united. In fact, the theme may define boundary or test of relevancy for stories being entered for a specific exhibit or round of research. The theme is comprised of a name and a description. The description is where the body of the theme exists and should describe high level changes and movements, while avoiding detailed descriptions of the entities being referenced.

Media. The media entity is designed to provide evidential and media support for other entities within the database. It is the only entity that can hold non-textual data. The ‘image’ property may be utilized for photos, scanned newspaper articles, advertisements, drawings or any type of static visual. The ‘audio’ property may contain music, audio testimonies, and recordings of events or other audio data. Finally, the ‘movie’ property could be used to hold video recordings of events, video testimonies, television programs or commercials, animations or other video data. It is generally assumed that only one of these fields will be used at a time, but in theory, all three could contain data simultaneously. The name property of the media should contain the media format followed by the media content (e.g. “Audio Recording of Sue Stein”, “Photo of Bill Postner”, “Video of Independence Parade.”) All non-textual media elements will be constrained to open standard formats: JPEG or TIFF for photos, MP3 or AIFF for audio, and H.264 or AVI for video files. It should be noted that multiple Media entities may be associated with other entities via linking.

Links Between Entities

In addition to the properties that belong to each type of entity, every entity also contains an implicit property which is a list of links to other entities. Any entity can be linked to any other entity, and any entity can

have any number of links. For example, two people can be linked together via a single link which describes their relationship. Likewise, an institution can be linked to a location, so that the Eldridge Street Synagogue (the institution) can be linked to the location that is its address. Similarly, an event, such as “the Kosher Meat Boycott”, may be linked to any number of people, including Isaac Gellis, and to the theme “The Role of Women.” All links are bidirectional, meaning that when A is linked to B, B is automatically linked back A.

While the relationship between two entities that are linked together is occasionally self-evident (such as Eldridge Street to its address), this is not always the case. For example, the relationship between two men could be as brothers or as business partners or one could be the other’s son. To make the relationship defined by a link clear, every link has an optional description as shown in the following table.

Like the descriptions associated with all of the entities, the description attached to a link may be brief or may be quite lengthy. In fact, it is not unusual that the heart of a story lies in the connections between its characters. It is possible to envision an instance of the database where the majority of the content actually lives on the links between all of the entities.

For the sake of standardization, simplicity and convenience, the actual interface to the database may offer optional, pre-defined link descriptions between certain types of entities. Specifically, we see this being useful in the representation of familial relationships, which are not only common, but are also best described in a standard format. Pre-defined link descriptions may be useful when linking media entities to non-media entities as well.

Custom Entities

When working in domains where a similar kind of object appears very frequently, it is useful to have a way describing those special kinds of objects in a way that allows for direct comparison between them. For example, if one’s research is focused on the history of various schools within a community, one may wish to include the size of the school and its grade range as specially noted properties. For this situation, we have left open the option of defining custom entities.

A custom entity is an extension of one of the seven basic entities, but with additional custom properties. For example, a “school” could be an “institution” entity with the two additional custom properties: “number of students” and “grade range.” By making these properties explicit, it opens the possibility of searching and processing this data by future researchers. It also makes clear the importance of that specific kind of data in the analysis. Additionally, having explicit custom properties allows for that information to be presented in a special way in an interactive interface or website.

In order to prevent the proliferation custom entities, which would dilute the rigor of the database’s organization, the definition of custom entities is limited to those who possess administrator level access. Once a custom entity prototype is defined by an administrator, any user working with that database would have the option of creating entities of that type. Following the example from above, once an administrator defines the “school” entity prototype, all the other users will be able to create “school” entities themselves.

Eldridge Street Sample Profiles

From Concept to Practice

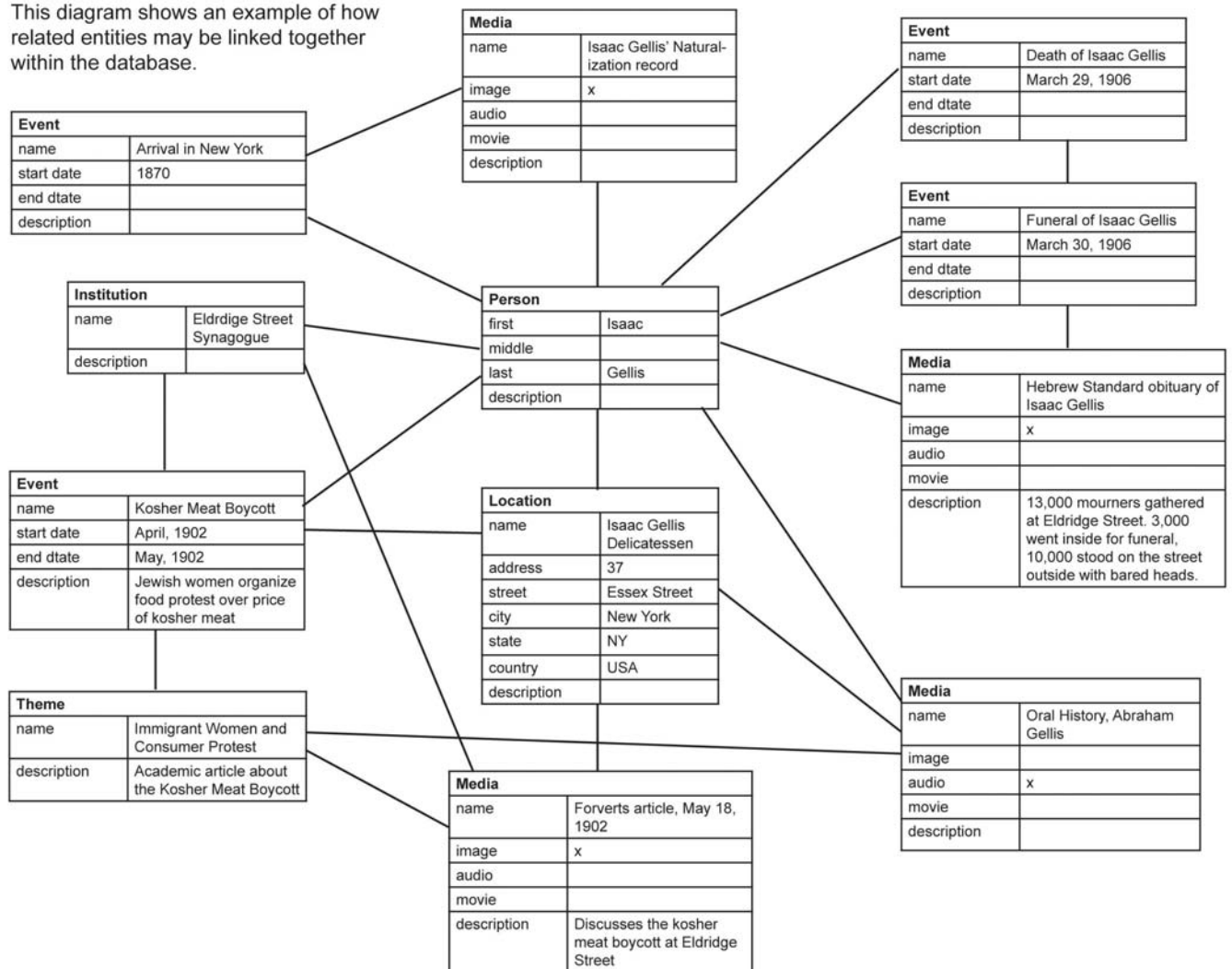
To research individuals from the Eldridge Street Synagogue’s historic congregation, names were picked from an existing list of congregants culled from minute books. We attempted to present a wide range of individuals, from an original president to a female entrepreneur who worshipped at Eldridge Street. The individuals became test cases for the database, allowing us to see whether the system we had created could adequately contain the disparate pieces of their lives and community. Sources for each individual were discovered by searching through internal Eldridge Street documents, such as minute books, constitutions and oral histories as

well as census records, naturalization records and newspaper archives. In doing the research, the archives of the New York Public Library, the New York Times and footnote.com proved fruitful and easily accessible digitally.

Once sources were gathered, the data was divided into discrete parts by entity. By coding each piece of information, which could be an entire story or simply one location, we tested whether the system we had created could encompass the varied pieces of data that had been collected. Changes were made to the database structure as information was added, such as space for citations and room for description and elaboration within an entity.

The diagram on this page shows an example of how related entities may be linked together within the database.

This diagram shows an example of how related entities may be linked together within the database.



Final Meeting Discussion

At the final meeting of the advisory group on February 16, there were several great discussions around the how the proposed database system could be used and what types of stories it could tell. Those discussions are captured here.

Open Systems

The first discussion centered around the notion of open systems. Specifically, there was talk of the need to make this system open from a technological perspective so that it could be modified by others and so that it could be incorporated into other open systems, such as Omeka. On a related note, the advisors also mentioned that it would be helpful to be able to link to external data such as the census database or data at Footnote.com, rather than trying to bring that data into this system.

Another half of the "open" conversation was in regards to user contributions. At the core of this discussion was the relationship of contributions by academics or the institution itself to contributions by average users. There was agreement that so-called "average" users had an enormous amount of information to contribute due to their large numbers, but there was also concern about the validity of their data. The group came to a general consensus that as long as contributors names were attached to the data they contributed, this issue would resolve itself. Josh Greenberg of NYPL presented several examples where contributions were tagged according to their source, so it was very clear what data was produced by the institution or by academics and what was not.

One powerful idea that also came up during this conversation was the notion that database contributors should themselves be part the database. This is especially relevant for Eldridge Street where all of the contributors are an essential part of the living history of the synagogue. That is, if one followed Eldridge's history from formation to the present, one would arrive at the group of individuals currently leading the institution and even the individuals visiting the institution. It also allows contributors to connect themselves to relatives and other individuals from the past who are already included in the database.

Content & Form

The second chapter of conversation revolved around the type of content that the database could hold. After much discussion, there was some explication that the group was actually talking about several different types of databases at the same time without making distinctions between them. For example, several advisors were imagining the database to function as a library or museum catalog, while others were considering an academic clipping or statistical database, while others still were thinking of a content management system for a website. In reality, the proposed database could handle almost all of these tasks, but in order to be successful with each, the institution would have to make this utilization scenario clear to the systems users. In other words, the institution would need to establish a priori what type of data the database should be holding and what type of stories they want to be telling, before entering the very first parcel of data.

By the end of the conversation, the advisory group arrived at a better sense of the type of content for which the proposed database was best suited. Susan Malbin presented the idea of Eldridge Street as a "house museum", which was a concept that seemed to resonate. As a house museum, Eldridge was interested in stories of individuals and their relationships to each other and their environment. In this case, the environment is both the Synagogue and the Lower East Side. Susan also mentioned the field of prosopography as a natural fit for the way this database is structured. Essentially, by collecting large amount of data, and looking at the connections and patterns between them, one can begin to understand general properties of the historical period in a local context.

Relationships

A major difference between the proposed database and traditional databases is the ability to link together any two entities and establish a connection or relationship between them. While this operation will be straightforward, establishing the characteristics and properties of these relationships is not so simple. This is especially true for relationships between individuals. Josh Greenberg of NYPL presented a website about the Yaddo conference in which all types of human relationships were shown for connecting two individuals. The advisory group then discussed several ambiguous cases. For example, is it worth establishing that two men are

brothers if they are both also connected as children of the same parents? Likewise, is it necessary or appropriate to mention casual relationships such as friendships or neighbors or romantic liaisons? Furthermore, how does one describe these relationships so that they may be accessed consistently at a later point in time. There was consensus among the group that the database interface would allow contributors to type in any relationship, but that the text field might auto-complete with pre-existing relationships so that the community would naturally define its own canon of relationship types. While there was some headway made during this part of the conversation, it was clear that the topic would need further refinement once the database was up and running with actual contributors.

Next Steps

The Museum at Eldridge Street hopes to secure funding to develop an interactive prototype (whether that be an interactive website, on-site history table, combination of the two, or new hardware) based on the structure and type of data reflected in this white paper. Our prototype will be field-tested with Museum visitors (either to our website and/or to our facility) by a broad range of visitors, of all ages and backgrounds. Our objective would be to enable visitors to “experience” the lives of our site’s early congregants and their Lower East Side immigrant neighborhood through the individuals and businesses that existed at the turn of the 20th century. The data collected will feature photographs, naturalization records, archival film footage, newspaper articles, and oral histories, among other interactive and educational resources.

To enable this to happen, the next step in the technical process will be to assemble a database development team that can identify existing software, use an existing platform (such as Omeka) or develop new software to implement the database schema that has been established in this whitepaper. Ideally, this development should take place in such a way that the database can be deployed easily in multiple locations at various institutions. After the database has been developed, Eldridge Street would begin its process of research and content development. Eldridge Street will require in-person training where the developers walk the staff at the museum through a standard process of development. Documentation for deployment and usage of the software will also be produced.